

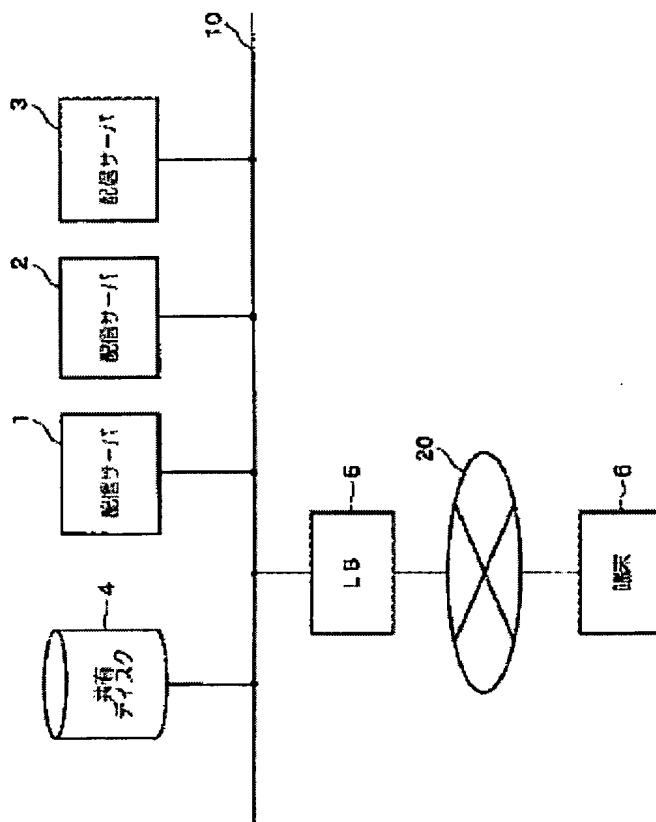
LOAD BALANCE CONTROL DEVICE AND LOAD BALANCE CONTROL METHOD

Patent number: JP2003256392
Publication date: 2003-09-12
Inventor: HIRANO MITSUNORI
Applicant: NTT COMWARE CORP
Classification:
- **International:** G06F15/177; G06F12/00; G06F13/00
- **European:**
Application number: JP20020057454 20020304
Priority number(s): JP20020057454 20020304

Report a data error here

Abstract of JP2003256392

<P>PROBLEM TO BE SOLVED: To provide a load balance control device capable of performing load balance at higher speed with less resource. **<P>SOLUTION:** The load balance control device 5 for distributing a load on multiple servers is provided with: a remainder calculation means for getting a remainder mod (n, M) to be calculated with the number M of multiple servers (1 to 3) set as a modulus to a numerical value n when a name of the content requested to distribute is regarded as the N-ary numerical value n where a number N is the number of kinds of characters used to constitute a name for specifying the content to distribute; and a distribution control means for letting the server associated with the remainder to be calculated by the remainder calculation means store the content corresponding to the numerical value n and distribute the content to clients requesting its distribution. **<P>COPYRIGHT:** (C) 2003,JPO



Data supplied from the **esp@cenet** database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-256392

(P2003-256392A)

(43) 公開日 平成15年9月12日 (2003.9.12)

(51) Int.Cl. ⁷	識別記号	F I	テマコード* (参考)
G 0 6 F 15/177	6 7 4	G 0 6 F 15/177	6 7 4 A 5 B 0 4 5
	6 7 8		6 7 8 B 5 B 0 8 2
12/00	5 4 6	12/00	5 4 6 A 5 B 0 8 9
13/00	3 5 7	13/00	3 5 7 Z
	5 2 0		5 2 0 C
審査請求 有 請求項の数 8 O L (全 8 頁)			

(21) 出願番号 特願2002-57454 (P2002-57454)

(22) 出願日 平成14年3月4日 (2002.3.4)

(71) 出願人 397065480

エヌ・ティ・ティ・コムウェア株式会社
東京都港区港南一丁目9番1号

(72) 発明者 平野 光徳

東京都港区港南一丁目9番1号 エヌ・テ
ィ・ティ・コムウェア株式会社内

(74) 代理人 100064908

弁理士 志賀 正武 (外2名)

Fターム(参考) 5B045 BB28 BB42 GG02 JJ07 JJ13

JJ26

5B082 HA05 HA08

5B089 GA04 GB09 JA33 KA06 KB07

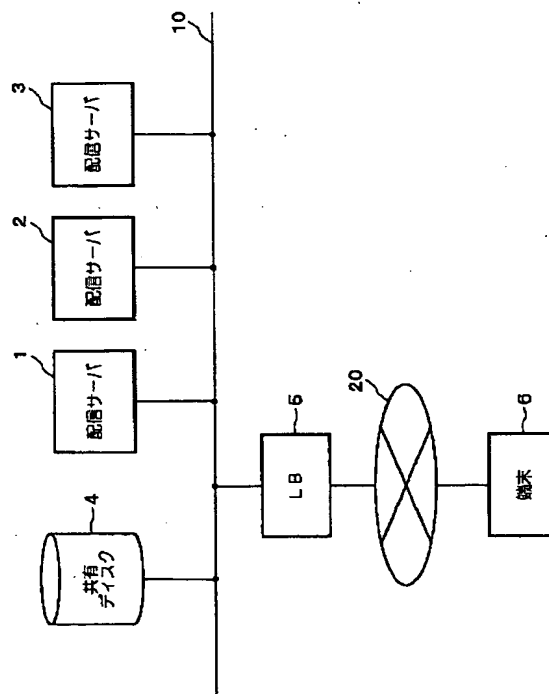
MA03

(54) 【発明の名称】 ロードバランス制御装置及びロードバランス制御方法

(57) 【要約】

【課題】 より少ない資源でより高速にロードバランスを実行できるロードバランス制御装置を提供する。

【解決手段】 複数のサーバにかかる負荷を分散させるためのロードバランス制御装置 (5) に、配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数をNとして、配信要求されたコンテンツの名称をN進数の数値とみなした場合の数値がnである場合、このnに対し複数のサーバ (1~3) の台数Mを法として算出される剰余mod (n, M) を求める剰余算出手段と、剰余算出手段により算出される剰余に対応させたサーバにnに対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向けこのコンテンツを配信させる配信制御手段とを備える。



【特許請求の範囲】

【請求項1】 複数のサーバにかかる負荷を分散させるためのロードバランス制御装置において、

配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記 n に対し前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手段と、

前記剰余算出手段により算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手段と、を具備することを特徴とするロードバランス制御装置。

【請求項2】 前記ロードバランス制御装置は、前記サーバによるコンテンツの配信が可能か否かを判断する診断手段をさらに備え、

前記剰余算出手段は、算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手段により配信が不可能であると診断された場合には、新たに剰余 $\text{mod}(n+1, M)$ を算出し、

前記配信手段は、前記剰余算出手段により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させることを特徴とする請求項1に記載のロードバランス制御装置。

【請求項3】 複数のサーバにかかる負荷を分散させるためのロードバランス制御方法において、

配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手段と、

前記剰余算出手段で算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手段と、を含むことを特徴とするロードバランス制御方法。

【請求項4】 前記ロードバランス制御方法は、前記サーバによるコンテンツの配信が可能か否かを判断する診断手順をさらに含み、

前記剰余算出手段が算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手順により配信が不可能であると診断された場合には、前記剰余算出手段において新たに剰余 $\text{mod}(n+1, M)$ を算出し、

前記配信手順においては、前記剰余算出手段により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させるこ

とを特徴とする請求項3に記載のロードバランス制御方法。

【請求項5】 複数のサーバにかかる負荷を分散させるためのロードバランス制御装置を制御するプログラムであって、

配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手順と、

前記剰余算出手順で算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手順と、を含むことを特徴とするプログラム。

【請求項6】 前記プログラムは、

前記サーバによるコンテンツの配信が可能か否かを判断する診断手順をさらに含み、

前記剰余算出手順が算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手順により配信が不可能であると診断された場合には、前記剰余算出手順において新たに剰余 $\text{mod}(n+1, M)$ を算出し、

前記配信手順においては、前記剰余算出手段により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させることを特徴とする請求項5に記載のプログラム。

【請求項7】 コンピュータにインストールされることにより、そのコンピュータが請求項1に記載のロードバランス制御装置となるプログラムを記録した記録媒体であって、

配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手順と、

前記剰余算出手順で算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手段と、を記録したコンピュータ読み取り可能な記録媒体。

【請求項8】 請求項7に記載のコンピュータ読み取り可能な記録媒体において、

前記サーバによるコンテンツの配信が可能か否かを判断する診断手順をさらに記録し、

前記剰余算出手順が算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手順により配信が不可能であると診断された場合には、前記剰余算出手順におい

て新たに剰余 $\text{mod}(n+1, M)$ を算出し、前記配信手順においては、前記剰余算出手順により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させるようにしたことを特徴とするコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、複数のサーバにかかる負荷を分散させるためのロードバランス制御装置に関する。

【0002】

【従来の技術】現在では、インターネット上の高速通信環境が整ってきており、ISP (Internet Service Provider) 等では、こうした高速通信環境に対応したコンテンツの提供（音楽や映像のストリーミング配信等）が行われるようになってきている。このような高速通信環境に対応して提供される音楽や映像のストリーミング配信等がなされる場合、1台の配信サーバにかかる負荷はさらに大きくなる。当然のことながら1台の配信サーバが配信できる能力には限界があり、一定以上の配信を行うことはできない。そのため、複数台の配信サーバでシステムを構成するとともに、ロードバランサ（ロードバランス制御装置）により各配信サーバにかかる負荷を分散させることが行われている。

【0003】

【発明が解決しようとする課題】しかしながら、従来のロードバランサでは、例えばラウンドロビン方式の場合、システムを構成する複数台の配信サーバに対して無作為に配信依頼を与える。そして、この配信依頼を受けた配信サーバは、要求されたコンテンツを配信するが、その記憶装置に必要なコンテンツを格納していない場合には、配信対象のコンテンツが蓄積されている共有ディスク等の他の場所から要求されたコンテンツを取り込むことになり、このため配信に係る処理能力が低下してしまう。したがって、同一のコンテンツは、無作為に選択された配信サーバからではなく、なるべく同一の配信サーバから配信する方が、他の場所から必要なコンテンツを取り込むための処理が削減され効率的となる。

【0004】そこで、従来のロードバランサでは、さらに、ユーザが指定したURLと過去に配信依頼を受けた配信サーバを識別する情報（IPアドレス等）を対応づけたリストを、ロードバランサ内に設けられたメモリ等の記憶媒体上に作成しておき、新たな配信依頼があった時には、ユーザが指定したURLとこのリストを比較して、このリストにユーザが指定したURLが存在する場合には、以前に配信依頼を出した配信サーバに配信依頼をする方式をとっていた。しかし、この方式の場合に

は、URLの文字列と配信サーバとを対応づけるリストを保存するための領域が必要であったため、かなりのメモリ容量が必要となっていた。また、ユーザが指定したURLとリストに登録されたURLとを比較する必要があるため、そのための処理時間を要していた。

【0005】本発明は、上記の点に鑑みてなされたもので、複数のサーバにかかる負荷を分散させるためのロードバランス制御装置において、より少ない資源でより高速にロードバランスを実行できるロードバランス制御装置を提供するものである。

【0006】

【課題を解決するための手段】請求項1に記載の発明のロードバランス制御装置は、複数のサーバにかかる負荷を分散させるためのロードバランス制御装置において、配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記 n に対し前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手段と、前記剰余算出手段により算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手段と、を具備することを特徴とする。

【0007】また、請求項2に記載の発明は、請求項1に記載のロードバランス制御装置において、前記サーバによるコンテンツの配信が可能か否かを判断する診断手段をさらに備え、前記剰余算出手段は、算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手段により配信が不可能であると診断された場合には、新たに剰余 $\text{mod}(n+1, M)$ を算出し、前記配信手段は、前記剰余算出手段により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させることを特徴とする。

【0008】また、請求項3に記載の発明のロードバランス制御方法は、複数のサーバにかかる負荷を分散させるためのロードバランス制御方法において、配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手段と、前記剰余算出手段で算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手順と、を含むことを特徴とする。

【0009】また、請求項4に記載の発明は、請求項3に記載のロードバランス制御方法において、前記サーバ

によるコンテンツの配信が可能か否かを判断する診断手順をさらに含み、前記剰余算出手順が算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手順により配信が不可能であると診断された場合には、前記剰余算出手順において新たに剰余 $\text{mod}(n+1, M)$ を算出し、前記配信手順においては、前記剰余算出手順により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させることを特徴とする。

【0010】また、請求項5に記載の発明は、複数のサーバにかかる負荷を分散させるためのロードバランス制御装置を制御するプログラムであって、配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手順と、前記剰余算出手順で算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手順と、を含むことを特徴とする。

【0011】また、請求項6に記載の発明は、請求項5に記載の発明において、前記サーバによるコンテンツの配信が可能か否かを判断する診断手順をさらに含み、前記剰余算出手順が算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手順により配信が不可能であると診断された場合には、前記剰余算出手順において新たに剰余 $\text{mod}(n+1, M)$ を算出し、前記配信手順においては、前記剰余算出手順により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させることを特徴とする。

【0012】また、請求項7に記載の発明は、コンピュータにインストールされることにより、そのコンピュータが請求項1に記載のロードバランス制御装置となるプログラムを記録した記録媒体であって、配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求める剰余算出手順と、前記剰余算出手順で算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させる配信制御手順と、を記録したコンピュータ読み取り可能な記録媒体である。

【0013】また、請求項8に記載の発明は、請求項7

に記載のコンピュータ読み取り可能な記録媒体において、前記サーバによるコンテンツの配信が可能か否かを判断する診断手順をさらに記録し、前記剰余算出手順が算出した前記剰余 $\text{mod}(n, M)$ に対応するサーバが、前記診断手順により配信が不可能であると診断された場合には、前記剰余算出手順において新たに剰余 $\text{mod}(n+1, M)$ を算出し、前記配信手順においては、前記剰余算出手順により算出される $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させるようにしたことを特徴とする。

【0014】

【発明の実施の形態】以下、本発明の実施の形態を、図面を参照して説明する。図1は、本発明の一実施の形態であるロードバランス制御装置を含むシステム構成を示すブロック図であり、図2は、同実施の形態のロードバランス制御装置の概略構成を示すブロック図である。

【0015】図1において、符号5は、本実施の形態のロードバランス制御装置(LB)を示し、このロードバランス制御装置5には、ローカルネットワーク10を介して、コンテンツを配信するための3台の配信サーバ(配信サーバ1~3)が接続されている。さらに、このローカルネットワーク10には、配信対象となっているコンテンツが蓄積される共有ディスク装置4が接続されている。また、符号6は、ユーザが使用する端末であり、インターネット20ならびにロードバランス制御装置5を介して、ユーザによる操作に応じた配信要求を発信し、配信サーバ1~3のいずれかから所望のコンテンツの配信を受ける。

【0016】ロードバランス制御装置5は、配信サーバ1~3が接続されるローカルネットワーク10およびインターネット20と接続するための入出力インターフェースである入出力部51と、後述する各処理を実行する処理部52と、本ロードバランス制御装置5を制御する制御プログラムならびに各種データを記憶する記憶部53とからなっている。このロードバランス制御装置5は、ユーザ側端末6から配信要求を受けると、配信サーバ1~3にかかる負荷を分散させるために配信サーバ1~3のいずれかを後述の方法で選択し、選択した配信サーバに対しユーザ側から要求されたコンテンツを端末6へ配信させるための配信命令を与える制御を行う。

【0017】一方、配信サーバ1~3は、ロードバランス制御装置5から受ける配信命令に応じて、ユーザ側から要求されたコンテンツのデータを、配信要求をしたユーザ側端末6に向け送出する。このとき、必要なコンテンツが配信命令が与えられた配信サーバに格納されている場合には、このコンテンツのデータを送出し、必要なコンテンツが当該配信サーバに格納されていない場合には、このコンテンツのデータを共有ディスク装置4から

取得して格納し、配信要求をしたユーザ側端末6に向け送出する。

【0018】ここで、ユーザ側の端末6から配信サーバ1～3へのアクセスは、WWW (World Wide Web) システムによるものとする。すなわち、端末6から出される配信要求は、WWWシステムにおけるURL (Uniform Resource Locator) の指定による。すなわち、この配信要求には、ユーザが端末6において入力/指定した、配信を希望するコンテンツを指定するURLが含まれており、このURLには、さらにコンテンツのファイル名が含まれている。また、URLに含まれるドメイン名は、別途DNSサーバ (図示せず) により解決されるものとする。

【0019】なお、ローカルネットワーク10におけるデータ通信は、ここではTCP/IPによるものとする。また、処理部52はメモリおよびCPU (中央処理装置) 等により構成され、処理部52の各機能を実現するためのプログラム (図示せず) をメモリにロードして実行することによりその機能が実現されるものとする。また、記憶部53は、ROM (Read Only Memory)、ハードディスク、光磁気ディスク等の不揮発性の記録装置により構成されている。

【0020】次に、このように構成された本実施形態のロードバランス制御装置5において実行される処理ならびに配信サーバ1～3の動作の詳細について説明する。

【0021】以下では、配信対象のコンテンツを指定する際用いられる名称 (ここでは、拡張子を除くファイル名) を構成するために使用される文字の種類数をNとし、配信要求されたコンテンツの名称をN進数の数値とみなした場合の数値をnとし、使用する配信サーバの台数をMとする。また、nに対しMを法として算出される剰余を $\text{mod}(n, M)$ で表す。すなわち、 $\text{mod}(n, M)$ は、nをMで除算した場合の余り (剰余) である。

【0022】例えば、上記名称を構成するために使用する文字を0～9の数字と、アルファベットの大文字であるA～Zを使用する場合、文字の種類数Nは36となる。この場合、0～9の数字と、アルファベットの大文字であるA～Zからなるファイル名 (拡張子を除く) を36進数の数値 (順序関係は、 $0 < 1 < \dots < 9 < A < B < \dots < Z$ とする) として処理する。具体的に、ファイル名が「ABA2」で、配信サーバの台数Mが3である場合 (例えば、図1の場合)、剰余 $\text{mod}(ABA2, 3)$ は、 $A \times 36^3 + B \times 36^2 + A \times 36 + 2$ を3で除算した余りである2となる。

【0023】なお、上記使用される文字の種類数N (=36) および配信サーバの台数M (=3) ならびに各剰余の値 (0～2) と配信サーバ1～3を対応づける情報は、予め、ロードバランス制御装置5の記憶部53に設定されているものとする。また、上記各剰余の値

(0～2) と配信サーバ1～3を対応づける情報としては、図3に示す配信制御テーブル等を用いることができる。同図に示す配信制御テーブルでは、各剰余の値に対して、各配信サーバのローカルIPアドレスを一意に割り当てている。

【0024】以下では、図4を参照し、(A) 全配信サーバが正常に動作している場合と、(B) いずれかの配信サーバが配信不能となっている場合に分けて説明する。

【0025】(A) 全配信サーバが正常に動作している場合：はじめに、ロードバランス制御装置5では、ユーザ側端末6から配信要求を受けると (ステップS11)、この配信要求からコンテンツのファイル名 (ここでは、「ABA2」とする) を抽出する (ステップS12)。次に、このファイル名 (「ABA2」) を上記のようにN進数の数値nとみなし (ここでは、 $n = ABA2$)、このnに対しM (=3) を法として算出される剰余 $\text{mod}(n, M) = \text{mod}(ABA2, 3)$ を求める (ステップS13)。ここでは、剰余 $\text{mod}(ABA2, 3)$ として2が得られる。

【0026】そして、算出された剰余2に対応する配信サーバ (ここでは、配信サーバ2が対応するものとする) が配信可能か診断する (ステップS14)。ここで、別途定められる手続を実行しこの手続に対し配信サーバ2が正常に応答するか、あるいは、応答しないか配信不能を示す所定のエラーメッセージを返すかで、配信サーバ2が配信可能か配信不能かを診断する。ここでは、全配信サーバが正常に動作しているので、配信サーバ2は正常に応答する。したがって、ステップS15では、配信可能であると判断し (OKの判定)、配信サーバ2に対し、当該ファイル名のコンテンツをユーザ側端末6に配信させるための配信命令を出す (ステップS16)。なお、この配信命令には、ユーザ側端末6のIPアドレスと配信対象のコンテンツのファイル名が含まれている。

【0027】次に、配信サーバ2は、ロードバランス制御装置5から上記配信命令を受けると (ステップS21)、この配信命令に含まれるコンテンツのファイル名から、当該コンテンツが配信サーバ2に格納されているか検索する (ステップS22)。そして、当該コンテンツが配信サーバ2に格納されている場合 (ステップS22の判断で、「あり」の判定の場合)、このコンテンツを配信命令に含まれるユーザ側端末6のIPアドレスを基に端末6に向け配信する (ステップS25)。

【0028】一方、ユーザから要求されたコンテンツが配信サーバ2に存在しない場合 (ステップS23の判断で、「なし」の判定の場合)、共有ディスク装置4から、当該コンテンツのデータを取得し記憶するとともに (ステップS24)、このコンテンツをユーザ側端末6に向け配信する (ステップS25)。以上のようにし

て、各配信サーバにかかる負荷が分散されると同時に、同じコンテンツは同じ配信サーバから配信されるようになる。

【0029】(B) いずれかの配信サーバが配信不能となっている場合：ここでは、配信サーバ1、3は正常に動作しているが、配信サーバ2が故障や配信能力を超える状況にあって配信不能となっているものとする。

【0030】はじめに、ロードバランス制御装置5では、ユーザ側端末6から配信要求を受けると(ステップS11)、この配信要求からコンテンツのファイル名(ここでは、「ABA2」とする)を抽出する(ステップS12)。次に、このファイル名(「ABA2」)を上記のようにN進数の数値nとみなし(ここでは、 $n = ABA2$)、このnに対し $M (= 3)$ を法として算出される剰余 $\text{mod}(n, M) = \text{mod}(ABA2, 3)$ を求める(ステップ13)。ここでは、剰余 $\text{mod}(ABA2, 3)$ として2が得られる。

【0031】そして、算出された剰余2に対応する配信サーバ(ここでは、配信サーバ2が対応するものとする)が配信可能か診断する(ステップS13)。ここでは、前述の(A)全配信サーバが正常に動作している場合と同様に診断するが、ここでは、上記剰余2に対応する配信サーバ2が配信不能であるので、ステップS15で、この配信サーバ2が配信不能であると判定され(NGの判定)、ステップS13に戻る。そして、ステップS13では、新たに剰余 $\text{mod}(n+1, M) = \text{mod}(ABA2+1, 3) = 0$ を算出する。

【0032】次に、新たに算出された剰余0に対応する配信サーバ(ここでは、配信サーバ3が対応するものとする)が配信可能か診断する(ステップS13)。ここでは、上記と同様に診断するが、ここでは、上記値3に対応する配信サーバ3が配信可能であるので、ステップS15で、この配信サーバ3が配信可能であると判定され(OKの判定)、ステップS16に移る。

【0033】そして、ステップS16で、配信サーバ3に対し、当該ファイル名のコンテンツをユーザ側端末6に配信させるための配信命令を出す。以降の配信サーバ3における処理は、前述の(A)全配信サーバが正常に動作している場合の配信サーバ2において行われる処理と同様に処理され、ユーザ側から要求されたコンテンツが端末6へ配信される。

【0034】以上のようにして、配信サーバ1~3のいずれかが配信不能となった場合であっても、同じコンテンツは同じ配信サーバから配信されるようになり、無作為に配信サーバが選択される場合よりも効率の良いロードバランスを実現できる。なお、上記例では、配信サーバ3が配信サーバ2を肩代わりすることになり、記憶するコンテンツの量が多くなるが、記憶容量を超える場合には、例えば、一般的なLRU(Least Recently Used)方式により、必要に応じて最もア

クセス頻度が少ないコンテンツから削除することで対応できる。なお、上記で説明した各動作フローは一例であり、上記の処理の流れに限定されるものではない。また、ステップS13における処理は、剰余算出手順に対応し、ステップS16における処理は、配信制御手順に対応し、ステップS14、S15における処理は、診断手順に対応する。

【0035】以上、詳細に説明したように、本実施の形態のロードバランス制御装置5では、従来のように、アクセスされたコンテンツのURLを保存する必要がなく、また、前述のように新規にアクセスされたコンテンツのURLと保存されたリスト中のURLを比較する必要がないため、高速に少量のメモリにて効率的なロードバランスを実現することができる。また、数字のみのファイル名のように、コンテンツのファイル名をN進数の数値として扱うので、古いコンテンツから順番に順序づけられるようにファイル名を付与した場合には、新しいコンテンツほど配信要求が出される可能性が高いことから、より均等に配信サーバにロードバランスされることになる。

【0036】なお、本発明は、インターネットの他、LANやダイヤルアップによるネットワークを利用してもよい。また、図1における処理部の機能の一部あるいは全部を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによりロードバランスの制御処理を行ってもよい。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。

【0037】また、「コンピュータシステム」は、WWWシステムを利用している場合であれば、ホームページ提供環境(あるいは表示環境)も含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムが送信された場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリ(RAM)のように、一定時間プログラムを保持しているものも含むものとする。

【0038】また、上記プログラムは、このプログラムを記憶装置等に格納したコンピュータシステムから、伝送媒体を介して、あるいは、伝送媒体中の伝送波により他のコンピュータシステムに伝送されてもよい。ここで、プログラムを伝送する「伝送媒体」は、インターネット等のネットワーク(通信網)や電話回線等の通信回線(通信線)のように情報を伝送する機能を有する媒体

のことをいう。また、上記プログラムは、前述した機能の一部を実現するためのものであっても良い。さらに、前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるもの、いわゆる差分ファイル（差分プログラム）であっても良い。

【0039】以上、この発明の実施形態を、図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の構成等も含まれる。

【0040】

【発明の効果】以上、詳細に説明したように、本発明によれば、配信対象のコンテンツを指定する際用いられる名称を構成するために使用される文字の種類数を N として、配信要求されたコンテンツの名称を N 進数の数値とみなした場合の数値が n である場合、前記複数のサーバの台数 M を法として算出される剰余 $\text{mod}(n, M)$ を求め、前記剰余算出手順で算出される剰余 $\text{mod}(n, M)$ に対応させたサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させるので、従来のように、アクセスされたコンテンツのURLを保存する必要がなく、また、前述のように新規にアクセスされたコンテンツのURLと保存されたリスト中のURLを比較する必要がないため、高速に少量のメモリにて効率的なロードバランスを実現することができる。また、同じコン

テンツは同じ配信サーバから配信されるようになり、無作為に配信サーバが選択される場合よりも効率の良いロードバランスを実現できる。

【0041】また、本発明によれば、算出された剰余 $\text{mod}(n, M)$ に対応するサーバが、配信不能であると診断された場合には、新たに剰余 $\text{mod}(n+1, M)$ を算出し、算出された $\text{mod}(n+1, M)$ に対応するサーバに前記 n に対応するコンテンツを記憶させるとともに、配信を要求したクライアントに向け該コンテンツを配信させるので、複数のサーバの内、いずれかが配信不能の状態となっても、上記効果を得ることができる。

【図面の簡単な説明】

【図1】 本発明の一実施の形態であるロードバランス制御装置を含む全体構成の概略図である。

【図2】 同実施の形態のロードバランス制御装置の概略構成を示す図である。

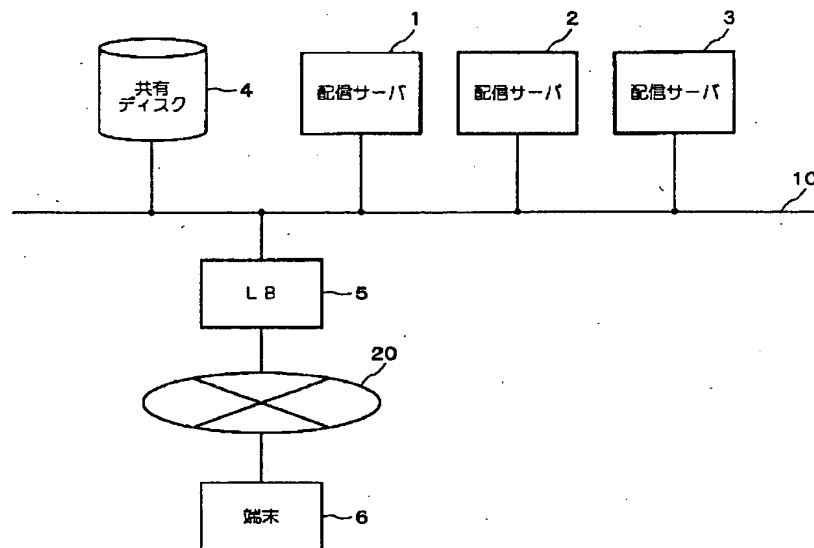
【図3】 配信制御テーブルの一例である。

【図4】 同実施の形態のロードバランス制御装置の動作を説明するための図である。

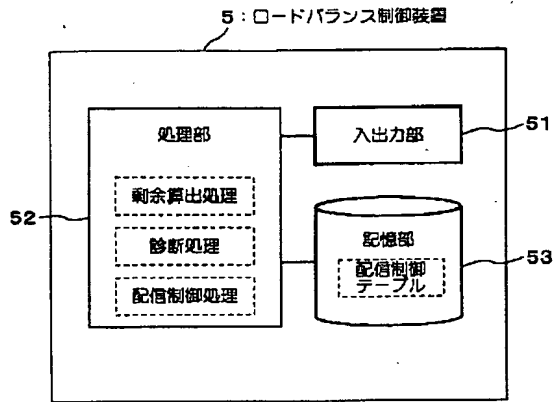
【符号の説明】

- 1, 2, 3…配信サーバ（サーバ）
- 4…共有ディスク装置
- 5…ロードバランス制御装置（LB）
- 6…端末
- 10…ローカルネットワーク
- 20…インターネット

【図1】



【図2】



【図3】

剰余	配信サーバ・アドレス
1	配信サーバ1のローカルIPアドレス
2	配信サーバ2のローカルIPアドレス
0	配信サーバ3のローカルIPアドレス

【図4】

